

# EVALUATING THE CAUSAL EFFECT OF UNIVERSITY GRANTS ON STUDENT DROPOUT: EVIDENCE FROM A REGRESSION DISCONTINUITY DESIGN USING PRINCIPAL STRATIFICATION

BY FAN LI<sup>†,‡</sup>, ALESSANDRA MATTEI<sup>\*,§</sup> AND FABRIZIA MEALLI<sup>§</sup>

*Duke University<sup>‡</sup> and University of Florence<sup>§</sup>*

Regression discontinuity (RD) designs are often interpreted as local randomized experiments: a RD design can be considered as a randomized experiment for units with a realized value of a so-called forcing variable falling around a pre-fixed threshold. Motivated by the evaluation of Italian university grants, we consider a fuzzy RD design where the receipt of the treatment is based on both eligibility criteria and a voluntary application status. Resting on the fact that grant application and grant receipt statuses are post-assignment (post-eligibility) intermediate variables, we use the principal stratification framework to define causal estimands within the Rubin Causal Model. We propose a probabilistic formulation of the assignment mechanism underlying RD designs, by re-formulating the Stable Unit Treatment Value Assumption (SUTVA) and making an explicit local overlap assumption for a subpopulation around the threshold. A local randomization assumption is invoked instead of more standard continuity assumptions. We also develop a model-based Bayesian approach to select the target subpopulation(s) with adjustment for multiple comparisons, and to draw inference for the target causal estimands in this framework. Applying the method to the data from two Italian universities, we find evidence that university grants are effective in preventing students from low-income families from dropping out of higher education.

**1. Introduction.** Amid the recent economic crisis in Europe, there has been a heated debate on how to arrange college students financial support, especially in terms of the instruments used, e.g., loans, grants, tuition waiver. Accurate evaluation of the effectiveness of the existing financial aid systems is crucial for providing information to policy makers to choose between different instruments. In Italy state universities offer financial aid every year to a limited number of eligible freshmen. The main objective of this intervention is to give equal opportunity to achieve higher education to motivated students irrespective of their economic background. Dropout from university is a relevant phenomenon in Italy: indeed, the low rate of university graduates among Italian youths is mainly due to the high dropout rate (about 30%) rather than to a low enrollment rate. In this paper, we will investi-

---

\*Part of this paper was written when Alessandra Mattei was a senior fellow of the Uncertainty Quantification program of the U.S. Statistical and Applied Mathematical Sciences Institute (SAMSI).

<sup>†</sup>Supported by U.S. NSF grants SES-1155697 and SES-1424688

*Keywords and phrases:* Bayesian, Causal effects, Intermediate variables, Principal stratification, Randomization, Regression discontinuity, University grants

gate the causal effects of Italian university grants on preventing students from low-income families from dropping out of higher education, using data on first-year enrollees from two state universities.

In the Italian university system, only students who both meet a pre-fixed eligibility criteria and apply for a grant can receive the grant, consisting of tuition waiver, free meals and accommodation, and a limited amount of money around 3 000 Euros. The eligibility status depends on an economic measurement of the student's family income and assets falling below or above a pre-determined threshold. This allocation rule motivates us to adopt the regression discontinuity (RD) design framework for evaluation. RD design—a quasi-experimental design for causal inference—was first introduced in psychology by [Thistlethwaite and Campbell \(1960\)](#) and has become increasingly popular since the late 1990s in economics and other fields. Recent surveys can be found in [Cook \(2008\)](#); [Imbens and Lemieux \(2008\)](#); [van der Klaauw \(2008\)](#); [Lee and Lemieux \(2010\)](#). There are two general setups in RD designs, the sharp and the fuzzy RD designs. In the sharp RD design, the original form of the design, the treatment status is assumed to be a deterministic step function of a so-called *forcing variable* or *running variable*. All units with a realized value of the forcing variable on one side of a pre-fixed threshold are assigned to one regime and all units on the other side are assigned to the other regime. The basic idea underlying a RD analysis is that one can compare units with very similar values for the forcing variable, but different levels of treatment, to draw inference on the causal effect of the treatment at the threshold. Examples of sharp RD designs can be found, among others, in [Berk and de Leuw \(1999\)](#); [Lee \(2008\)](#); [Mealli and Rampichini \(2012\)](#). In the fuzzy RD design, the realized value of the forcing variable does not alone determine the receipt of the treatment, although a value of the forcing variable falling above or below the threshold acts as an encouragement or incentive to participate in the treatment. In those cases, the receipt of the treatment depends also on individual choices, which may confound treatment receipt. [Hahn, Todd and Van der Klaauw \(2001\)](#) establish a connection between fuzzy RD designs and the instrumental variable (IV) settings, and show that in a fuzzy RD setting one can identify the local average treatment effect ([Imbens and Angrist, 1994](#)) for a subpopulation of compliers at the threshold. Examples of fuzzy RD designs can be found, among others, in [van der Klaauw \(2002\)](#); [Battistin and Rettore \(2008\)](#); [Garibaldi et al. \(2012\)](#).

The Italian university grant allocation rule defines a fuzzy RD design because not all eligible students get a grant, they must apply first, and application is voluntary. Also ineligible students may apply, even if they will not receive any grant. Comparing to standard fuzzy RD designs where only assignment (eligibility) and receipt of the treatment (grant) are available, the additional data on the application status in this study can provide valuable information with important policy implications. In this article, we will show how to capitalize on application. In particular, a main methodological contribution of this article is to develop a framework for RD analysis that is fully consistent with the Rubin Causal Model (RCM, [Rubin, 1974, 1978](#)) using potential outcomes. Resting on the fact that grant

application and grant receipt statuses are post-assignment (post-eligibility) intermediate variables, we adopt the principal stratification framework (Frangakis and Rubin, 2002)—a generalization of the IV approach to noncompliance (Angrist, Imbens and Rubin, 1996; Imbens and Rubin, 1997)—to define causal estimands and lay the basis for inference.

In the literature, causal inference in RD designs is usually based on comparisons of units with close but distinct values of the forcing variable and relies on smoothness assumptions about the relationship between outcomes and the forcing variable around the threshold, which imply randomization at the single threshold value. For example, in fuzzy RDs, estimands are usually specified as ratio of differences of regression functions at the threshold, and inference generally relies on asymptotic approximations (e.g. Imbens and Lemieux, 2008). In real applications, large-sample approximations might be unreliable due to the small sample size, and exact inference would be preferable. RD designs have been often described as designs that lead to locally randomized experiments around the threshold (Lee, 2008; Lee and Lemieux, 2010; Dinardo and Lee, 2011). Expanding on this interpretation, a recent strand of the literature (e.g., Cattaneo, Frandsen and Titiunik, 2015; Sales and Hansen, 2014) is moving towards a formal and well-structured definition of the conditions under which RD designs are equivalent to local randomized experiments.

We further develop the idea of local randomization; our goal is to provide a formal definition of the hypothetical experiment underlying RD designs, based on a description of the assignment mechanism, i.e., the process that describes why some units got assigned to different treatments, formalized as a unit-exchangeable stochastic function of covariates and potential outcomes. The core of our framework is to assume there exists at least one subpopulation around the threshold where a *local overlap* assumption holds. For this subpopulation we explicitly introduce a *local randomization* assumption.

Though our framework is not tied to any mode of inference, we choose the Bayesian approach for reasons explained later. In particular, a second methodological contribution of this article lies in our development of a Bayesian hierarchical modeling approach to adjust for multiple comparisons in selecting the target subpopulation(s). Our work contributes to the limited literature on Bayesian analysis to RD (Chib and Jacobi, 2011; Chib and Greenberg, 2014), as well as to the literature on Bayesian causal inference (e.g., Rubin, 1978; Imbens and Rubin, 1997; Barnard et al., 2003; Elliott, Raghunathan and Li, 2010; Schwartz, Li and Mealli, 2011; Mattei, Li and Mealli, 2013).

In Section 2, we introduce the basic setup and the causal estimands. In Section 3, we propose a probabilistic formulation of the assignment mechanism for general RD designs, explicitly formulating the key assumptions, and elaborate it for the particular RD design used in the Italian university grants. Selection of the subpopulations where these assumptions hold is also discussed. A Bayesian approach for inferring causal effects in RD designs is developed in Section 4. We then apply the proposed approach to evaluate causal effects of Italian university grants on student dropout in Section 5. Section 6 concludes.

## 2. Causal estimands.

**2.1. Basic setup.** We introduce the notation in the context of Italian university grants. Let  $Z$  be the eligibility status, which is the initial assignment and plays the role of an “instrument” or an “encouragement” as in randomized experiments with noncompliance. Consider a sample or population of  $N$  units; each can be either eligible to receive a treatment,  $z = 1$ , or ineligible,  $z = 0$ . In the Italian grants system, eligibility depends on the value of a combined measurement of one’s assets including income and properties, adjusted for family size, denoted by  $S$ . If a student, satisfying preliminary grade criteria, has a value of  $S$  falling below a pre-determined threshold, e.g.  $s_0 = 15\,000$  euro, he/she is eligible, and not otherwise. That is, the eligibility status  $Z_i$  for student  $i$  is a deterministic function of  $S$ :  $Z_i = \mathbf{1}(S_i \leq s_0)$ , where  $\mathbf{1}(\cdot)$  is the indicator function. Using the terminology in RD designs,  $S$  is the forcing variable.

All variables measured after each unit  $i$  is assigned eligibility  $Z_i$ , namely, the application status, the receipt of the grant and the dropout status, are post-assignment variables, and, in principle, eligibility may affect them. Thus we can define potential outcomes for these variables: for each student  $i$  ( $i = 1, \dots, N$ ), given eligibility status  $z$  ( $z = 0, 1$ ), let  $A_i(z)$  be an indicator for the potential grant application status (equal to 1 if student  $i$  applies for a grant and 0 otherwise),  $W_i(z)$  be an indicator for the potential treatment received (equal to 1 if student  $i$  receives a grant and 0 otherwise), and  $Y_i(z)$  be the potential indicator for dropout (1 if student  $i$  drops out of university, 0 otherwise). These notations, with only two potential outcomes for each post-treatment variable for each unit, reflect the acceptance of the Stable Unit Treatment Value Assumption (SUTVA, [Rubin, 1980](#)), which implies that there is no interference between units and that there are no levels of the eligibility status other than zero and one. A more explicit formulation of SUTVA will be introduced in Section 3.1.

For each unit,  $i$ , given the observed eligibility status  $Z_i$ , the following variables are observed:  $A_i^{obs} = A_i(Z_i)$ , the observed application status;  $W_i^{obs} = W_i(Z_i)$ , the observed treatment received; and  $Y_i^{obs} = Y_i(Z_i)$ , the observed dropout status. The remaining potential outcomes are missing:  $A_i^{mis} = A_i(1 - Z_i)$ ,  $W_i^{mis} = W_i(1 - Z_i)$ , and  $Y_i^{mis} = Y_i(1 - Z_i)$ . A vector of  $p$  pre-treatment variables,  $\mathbf{X}_i$ , is also observed for each unit. We use boldface upper-case letters to denote the vector of values of a variable for all units from hereon. For example,  $\mathbf{Z} = (Z_1, \dots, Z_N)'$ ,  $\mathbf{A}^{obs} = (A_1^{obs}, \dots, A_N^{obs})'$ .

**2.2. The role of Principal Stratification for causal inference in Fuzzy RD designs.** In the RCM, a causal effect is defined as a comparison of the potential outcomes  $Y_i(1)$  and  $Y_i(0)$ , e.g.,  $\mathbb{E}[Y_i(1) - Y_i(0)]$ , for a common set of units. Obviously, in our study, such comparisons between potential dropout statuses only measure the effect of the eligibility status. To draw inference about the causal effect of *receiving a grant*, additional structure and assumptions are required. Since both the application status and receipt of the grant are post-assignment intermediate variables, we adopt the Principal Stratification

(Frangakis and Rubin, 2002) framework.

For each intermediate variable, principal stratification defines a cross-classification of subjects into groups, named principal strata, defined by the joint potential values of that intermediate variable under each of the assignments being compared. In our study, based on the application status  $A$ , subjects are classified into four (latent) principal strata,  $G_i \equiv (A_i(0), A_i(1))$ : compliant-applicants  $G_i = (0, 1) = CA$ , students who would not apply if ineligible, but would apply if eligible; always-applicants  $G_i = (1, 1) = AA$ , students who would apply irrespective of their eligibility status; never-applicants  $G_i = (0, 0) = NA$ , students who would not apply irrespective of their eligibility status; and defiant-applicants  $G_i = (1, 0) = DA$ , students who would not apply if eligible, but would apply if ineligible. Because principal strata are not affected by assignment, we can define *population-average* causal effects conditional on the principal strata, known as principal causal effects:

$$(1) \quad \tau_g \equiv \mathbb{E}[Y_i(1) - Y_i(0) | G_i = g],$$

for  $g = AA, CA, NA, DA$ . Then the average causal effect of eligibility on dropout is a weighted average of these principal causal effects:

$$\mathbb{E}[Y_i(1) - Y_i(0)] = \sum_{g=AA,CA,NA,DA} \pi_g \tau_g,$$

where  $\pi_g$  is the proportion of units in principal stratum  $g$ .

Never-applicants and defiant-applicants never receive a grant, so for them we always observe the outcome in the absence of the grant. By contrast, for always-applicants and compliant-applicants we can observe  $Y_i(1)$  for some eligible students who receive a grant and  $Y_i(0)$  for some other ineligible students who do not receive a grant. Therefore, always-applicants and compliant-applicants are the only groups where we can learn information about the effect of receiving a grant in this study, and thus the corresponding principal causal effects,  $\tau_{AA}$  and  $\tau_{CA}$ , are the causal estimands of primary interest.

In the standard IV approach to noncompliance (Angrist, Imbens and Rubin, 1996; Imbens and Rubin, 1997) as well as standard setting of fuzzy RD designs (e.g., Imbens and Lemieux, 2008), data on application status is not utilized, either because it is not available or because it is ignored. Instead, the analysis is based on the principal strata formed by the intermediate variable of grant receipt status. Specifically, there are four principal strata based on the joint potential grant receipt statuses,  $R_i = (W_i(0), W_i(1))$ : compliers  $R_i = (0, 1)$ , students who would receive the grant if eligible and would not receive the grant if ineligible; always-takers  $R_i = (1, 1)$ , student would receive the grant regardless of eligibility; never-takers  $R_i = (0, 0)$ , student would not receive the grant regardless of eligibility; and defiers  $R_i = (1, 0)$ , students who would not receive the grant if eligible and would receive the grant if ineligible. The focus is generally on the causal effect for compliers:

$$\tau \equiv \mathbb{E}[Y_i(1) - Y_i(0) | R_i = (0, 1)].$$

We now establish the connection between these two sets of principal strata. The Italian grant assignment rule implies that  $W_i(0) = 0$  for all  $i$ , as ineligible units have no access to a grant, and  $W_i(1) = 0$  if  $A_i(1) = 0$ , as eligible units need to apply for a grant to receive a grant. Therefore, by design, there are no always-takers or defiers, and the remaining principal strata  $R$ 's can be expressed as unions of principal strata  $G$ 's: never-takers comprise never-applicants and defiant-applicants, and compliers comprise always-applicants and compliant-applicants. As such,  $\tau$  can be rewritten as the weighted average of the causal effects for always-applicants and compliant-applicants:

$$(2) \quad \tau = \mathbb{E}[Y_i(1) - Y_i(0) \mid G_i \in \{AA, CA\}] = \frac{\pi_{AA}\tau_{AA} + \pi_{CA}\tau_{CA}}{\pi_{AA} + \pi_{CA}}.$$

This illustrates that principal strata defined by the application status leads to a finer partition of the units than principal strata defined by the grant-receipt status. Indeed the standard IV causal estimand—the causal effect for compliers  $\tau$ —provides information on a ‘marginal’ (with respect to application behavior) causal effect. If causal effects are homogeneous, marginalizing over application behavior does not critically affect the evaluation analysis. Conversely, if causal effects are heterogeneous, as we have found in this study, ignoring application behavior represents a loss of useful information with potentially important policy implications. For example, if the grants are found out to have a higher positive effect on always-applicants than on compliant-applicants, then it would be useful and cost-effective to study the characteristics of ineligible applicants and include those into the eligibility rule to allocate additional resources.

The estimands  $\tau_{AA}$ ,  $\tau_{CA}$  and  $\tau$  represent effects of eligibility, rather than effects of the receipt of a grant. However, “the receipt of a grant” is completely confounded with “the eligibility status”:  $W(z) = z \times A(z) = z$  for always-applicants and compliant-applicants. To *attribute* these effects to “the receipt of a grant”, below we can make an exclusion restriction assumption:

**ASSUMPTION 1.** (*Exclusion Restriction for Compliant-Applicants and Always-Applicants*). *For all units with  $G_i \in \{AA, CA\}$ , or equivalently  $R_i = (0, 1)$ , the effect of eligibility is only through the receipt of the grant.*

Assumption 1 attributes the intention-to-treat effect for compliers to the causal effect of the receipt of grant, rather than to its assignment (eligibility). A more formal version of this assumption, which requires double-indexed notations, is given in [Imbens and Rubin \(2015\)](#) (Chapter 23, Assumption 23.4). This type of exclusion restriction is routinely made, often implicitly, in randomized experiments with full compliance ([Mealli and Rubin, 2002](#); [Mealli and Pacini, 2013](#); [Imbens and Rubin, 2015](#)).

In real studies, the *sample-average* counterpart of the *population-average* estimands



may also be of interest:

$$(3) \quad \tau_g^S \equiv \frac{1}{N_g} \sum_{i:G_i=g} [Y_i(1) - Y_i(0)],$$

where  $g = AA, CA, \{AA, CA\}$  and  $N_g$  is the number of units in stratum  $g$ . Usually the *sample-average* effects can be estimated more precisely than their *population-average* counterparts. The subtle difference between them in Bayesian inference is explained in Section 4. More details can be found, for example, in [Rubin \(1978\)](#); [Imbens and Rubin \(1997\)](#) and [Imbens \(2004\)](#). For simplicity of notation, we do not make the distinction between population-average and sample-average estimands in the methodological discussion, but will present both estimates in the application.

### 3. The basis for inference.

**3.1. Probabilistic treatment assignment mechanism in RD designs.** The complex selection process in Italian university grants system implies that the mechanism governing the receipt of the grant, which depends on both institutional and individual choices, is not ignorable. Below we introduce a probabilistic assignment mechanism underlying the RD design considered here, which is also applicable to general RD settings with minor modifications.

We first define the assignment mechanism, which is a row-exchangeable function that assigns probabilities to all  $2^N$  possible  $N$ -dimensional vectors of assignments  $\mathbf{Z}$ , as a row-exchangeable function that assigns probabilities to all possible  $N$ -dimensional vectors of realizations of the forcing variable,  $\mathbf{S}$ , above or below the threshold value,  $s_0$ . Formally,

$$(4) \quad \begin{aligned} \Pr(\mathbf{Z} = \mathbf{z} | \mathbf{A}(0), \mathbf{A}(1), \mathbf{W}(0), \mathbf{W}(1), \mathbf{Y}(0), \mathbf{Y}(1), \mathbf{X}) \\ = \Pr(\mathbf{S} \in \Lambda | \mathbf{A}(0), \mathbf{A}(1), \mathbf{W}(0), \mathbf{W}(1), \mathbf{Y}(0), \mathbf{Y}(1), \mathbf{X}), \end{aligned}$$

where  $\mathbf{z} \in \{0, 1\}^N$  and  $\Lambda \in \{(-\infty, s_0]^N, (-\infty, s_0]^{N-1} \times (s_0, \infty), (s_0, \infty) \times (-\infty, s_0]^{N-1}, \dots, (-\infty, s_0] \times (s_0, \infty)^{N-1}, (-\infty, s_0]^{N-1} \times (s_0, \infty), (s_0, \infty)^N\}$ . Since  $Z$  is a deterministic function of  $S$ , the assignment mechanism can be formulated with respect to either  $Z$  or  $S$ . Here we prefer  $S$  because it is the underlying random variable that describes the reasons for the missing and observed values of potential outcomes: a value of  $S$  is assigned, which in turn determines a value for  $Z$ .

Statistical inference for causal effects requires assumptions on the assignment mechanism. We introduce assumptions that allow us to describe RD settings as classical randomized experiments around the threshold. The assignment mechanism in Equation (4) is a classical randomized experiment if (i) it is individualistic:

$$\begin{aligned} \Pr(\mathbf{S} \in \Lambda | \mathbf{A}(0), \mathbf{A}(1), \mathbf{W}(0), \mathbf{W}(1), \mathbf{Y}(0), \mathbf{Y}(1), \mathbf{X}) \\ = \prod_{i=1}^n \Pr(S_i \leq s_0 | A_i(0), A_i(1), W_i(0), W_i(1), Y_i(0), Y_i(1), \mathbf{X}_i); \end{aligned}$$

(ii) it is probabilistic, which implies that for each unit,  $i$ , both events  $S_i \leq s_0$  and  $S_i > s_0$  have *a priori* a non-zero probability of occurring; (iii) it is unconfounded, that is, free of dependence of any potential outcomes; and (iv) it is a known function of its arguments.

The particular assignment rules underlying RD designs suggest that these assumptions are more reasonable for subpopulations of units who have a relatively large probability that the realized values of the forcing variable fall in a neighborhood around the threshold,  $s_0$ . For these subpopulations, we can reasonably assume that the distribution of the forcing variable is unrelated to observed and unobserved characteristics of students. On the other hand students with a very small (close to zero) or a very large (close to one) probability that  $S_i \leq s_0$  are likely systematically different from other students. For example, potential outcomes observed for very rich students, who do not receive any grant, are plausibly different from potential outcomes for poor students with a realized value of  $S$  around the threshold, who do not receive a grant, and vice versa. Therefore we focus on subpopulations of students who have a probability that  $S_i \leq s_0$  strictly between zero and one, and sufficiently far away from zero and one. The following assumption guarantees that at least one such subpopulation of units exists.

**ASSUMPTION 2.** (*Local overlap*). *Let  $\mathcal{U}$  be the random sample (or population) of units in the study. There exists a subset of units,  $\mathcal{U}_{s_0}$ , such that for each  $i \in \mathcal{U}_{s_0}$ ,  $\Pr(S_i \leq s_0) > \epsilon$  and  $\Pr(S_i > s_0) > \epsilon$  for some sufficiently large  $\epsilon > 0$ .*

Assumption 2 assumes that there exists a subpopulation of units, each of whom has a non-zero probability of being assigned to either treatment levels. This represents a main distinction between our framework and the existing RD literature that often describes RD designs as settings where the overlap assumption is violated. Now within the subpopulation  $\mathcal{U}_{s_0}$  we can formally introduce a modified SUTVA specific to the RD settings:

**ASSUMPTION 3.** (*Local RD-SUTVA*). *For each  $i \in \mathcal{U}_{s_0}$ , consider two eligibility statuses  $Z'_i = \mathbf{1}(S'_i \leq s_0)$  and  $Z''_i = \mathbf{1}(S''_i \leq s_0)$ , with possibly  $S'_i \neq S''_i$ . If  $Z'_i = Z''_i$ , that is, if either  $S'_i \leq s_0$  and  $S''_i \leq s_0$ , or  $S'_i > s_0$  and  $S''_i > s_0$ , then  $A_i(\mathbf{Z}') = A_i(\mathbf{Z}'')$ ,  $W_i(\mathbf{Z}') = W_i(\mathbf{Z}'')$ , and  $Y_i(\mathbf{Z}') = Y_i(\mathbf{Z}'')$ .*

Local RD-SUTVA rules out interference between units, implying that potential outcomes for a student cannot be affected by the eligibility status of other students. Local RD-SUTVA also assumes that there are no levels of the eligibility status other than zero and one. This component of RD-SUTVA implies that values of the forcing variable leading to the same eligibility status cannot alter potential outcomes for any unit, and thus allows us to avoid defining potential outcomes as functions of the forcing variable. Under the local RD-SUTVA for each unit within  $\mathcal{U}_{s_0}$  there exist only two potential outcomes for each post-assignment variable, corresponding to the realized value of the forcing variable falling *below* and *above* the threshold, respectively.



Finally, we need to formalize the concept of RD design as local randomized experiment: in a neighborhood of the threshold  $s_0$  the forcing variable does not depend on either the potential outcomes or pre-treatment variables. Formally, we have:

ASSUMPTION 4. (*Local randomization*). For each  $i \in \mathcal{U}_{s_0}$ ,

$$\Pr(S_i | A_i(0), A_i(1), W_i(0), W_i(1), Y_i(0), Y_i(1), \mathbf{X}_i) = \Pr(S_i).$$

Assumption 4 states that within the subpopulation  $\mathcal{U}_{s_0}$  a Bernoulli trial has been conducted, with individual assignment probabilities, that is, the individual probabilities of being eligible to receive a grant, depending only on the distribution of the forcing variable:  $\Pr(Z_i = 1) = \Pr(S_i \leq s_0)$ . This assumption is crucial in justifying the key idea underlying any RD design. It implies that the eligibility statuses are randomly assigned in some small neighborhood,  $\mathcal{U}_{s_0}$ , around  $s_0$ .

Assumption 4 may not always be plausible. For instance, when the forcing variable is a deterministic variable, which conceptually cannot be interpreted as a random variable with a non-degenerate probability distribution (such as time), the underlying design cannot, in general, be interpreted as a local randomized experiment (see Section 6.3 in [Lee and Lemieux, 2010](#), pp 347).

There are subtle but substantive differences between local RD-SUTVA and local randomization. Local RD-SUTVA is an exclusion restriction assumption and it is required to make the representation of potential outcomes as functions of the eligibility status adequate. Local randomization is an independence assumption and it is crucial to make inference. RD-SUTVA is different from independence assumptions: it does not imply that the probability that we observe a value of the forcing variable above or below the threshold does not depend on potential outcomes. RD-SUTVA simply implies that the exposure to assignment level  $z$  specifies well-defined potential outcomes, for all unit  $i$  and assignment levels  $z$ . In other words, considering potential outcomes as random variables, RD-SUTVA does not imply that potential outcomes have the same distribution for each value of the forcing variable. In order to make the forcing variable independent of potential outcomes, we need to introduce additional assumptions, such as Assumption 4.

Following Assumption 2, we can define a local version of the target estimands:

$$(5) \quad \tau_{g,s_0} \equiv \mathbb{E}[Y_i(1) - Y_i(0) | G_i = g, i \in \mathcal{U}_{s_0}],$$

for  $g = AA, CA, \{AA, CA\}$  and their *finite-sample* counterparts, and we have:

$$\tau_{\{AA, CA\}, s_0} \equiv \tau_{s_0} = \frac{\tau_{AA, s_0} \pi_{AA, s_0} + \tau_{CA, s_0} \pi_{CA, s_0}}{\pi_{AA, s_0} + \pi_{CA, s_0}},$$

where  $\pi_{g, s_0} = \Pr(G_i = g | i \in \mathcal{U}_{s_0})$  for  $g = AA, CA, NA, DA$ , denote the proportion of principal strata in the subpopulation  $\mathcal{U}_{s_0}$ . A special case of  $\mathcal{U}_{s_0}$  contains the subpopulation

of units with a realized value of the forcing variable *exactly* equal to the threshold value,  $s_0$ .

It is worth noting that Assumption 4 implies that

$$\mathbb{E}[Y_i(1) - Y_i(0) | G_i = g, i \in \mathcal{U}_{s_0}] = \mathbb{E}[Y_i(1) - Y_i(0) | Z_i = 1, G_i = g, i \in \mathcal{U}_{s_0}].$$

Under the allocation rule of the Italian university grants,  $Z_i = W_i^{obs}$  for always-applicants and compliant-applicants. Therefore, the local randomization assumption allows the estimands  $\tau_{AA,s_0}$ ,  $\tau_{CA,s_0}$ , and  $\tau_{s_0}$  to be interpreted as causal effects of receiving a grant for subpopulations of students who actually receive a grant, analogous to the notion of average treatment effect for the treated.

**3.2. Two additional assumptions.** The following two assumptions—though not necessary for Bayesian inference—are plausible in our study and can sharpen the inference.

**ASSUMPTION 5.** *Monotonicity of Application Status:*

$$A_i(1) \geq A_i(0), \quad \text{for all } i \in \mathcal{U}_{s_0}.$$

**ASSUMPTION 6.** *Stochastic Exclusion Restriction for Never-Applicants:*

$$\Pr(Y_i(1)|G_i = NA, i \in \mathcal{U}_{s_0}) = \Pr(Y_i(0)|G_i = NA, i \in \mathcal{U}_{s_0}).$$

Monotonicity rules out the existence of defiant-applicants. The exclusion restriction rules out direct effects of eligibility on dropout for never-applicants. Never-applicants are students who would never apply for a grant irrespective of their eligibility status. These students would not receive the grant in any case. Exclusion restriction for never-applicants (Assumption 6) is of very different nature from the exclusion restriction for compliant-applicants and always-applicants (Assumption 1): Assumption 6 has implications for inference but not for interpretation, whereas Assumption 1 is made solely for interpreting the causal effects of assignment on the outcome attributable to the causal effects of treatment on the outcome. More discussions on the difference can be found in [Imbens and Rubin \(2015, Chapter 23\)](#) and [Mealli and Pacini \(2013\)](#).

**3.3. Selection of the subpopulations.** An important issue in practice is the selection of the subpopulation  $\mathcal{U}_{s_0}$  where the RD assumptions hold. There can be a diverse choice of the shape of the subpopulation. In this paper, we limit our choice to symmetric intervals with respect to  $s_0$ , for convenience and also to match the common practice of RD analysis. Specifically, we make the following assumption:

**ASSUMPTION 7.** *There exists  $h > 0$  such that for each  $\epsilon > 0$ ,  $\Pr(s_0 - h \leq S_i \leq s_0 + h) > 1 - \epsilon$ , for each  $i \in \mathcal{U}_{s_0}$ .*

Assumption 7 allows us to focus on the specific subsets of symmetric intervals among all neighborhoods of different shape around the threshold,  $s_0$ . Note that Assumptions 2 and 7 do not imply that  $\mathcal{U}_{s_0}$  is unique. They only require that there exists at least one subpopulation,  $\mathcal{U}_{s_0}$ . Consequently, we are not interested in finding the largest  $h$ , but we only aim at determining plausible values for  $h$ .

Our approach for selecting bandwidth  $h$  exploits the fact that Assumption 4 is a local randomization assumption, in the sense that it holds for a subset of units, but may not hold in general for other units. As such, under Assumption 4, in the subpopulation  $\mathcal{U}_{s_0}$ , pre-treatment variables should be well balanced in the two subsamples defined by assignment, and thus any test of the null hypothesis of no effect of assignment on pre-treatment covariates should fail to reject the null.

Assessing balance in the observed covariates raises problems of multiple comparisons, which may lead to a much higher than planned type I error if they are ignored (e.g., [Benjamini and Hochberg, 1995](#)). We account for multiplicities using a Bayesian hierarchical mixed model, which provides an explicit method for borrowing information across covariates (e.g., [Berry and Berry, 2004](#); [Scott and Berger, 2006](#)). Following [Berry and Berry \(2004\)](#), we use a mixture for the prior distribution of the eligibility parameters by assigning a point mass on equality of the means of the covariates between eligible and ineligible units. This Bayesian procedure provides a measure of the risk (posterior probability) that a chosen interval around the threshold,  $s_0$ , defines a subpopulation of units that does not exactly matches any true  $\mathcal{U}_{s_0}$ , including subjects for which our RD assumptions do not hold. More details are given in Section 5. The idea to exploit balance tests of pre-assignment variables to select a subpopulation of units is also used in [Cattaneo, Frandsen and Titiunik \(2015\)](#), but their approach aims at selecting the *largest* subpopulation and does not account for multiple comparisons.

Our approach parallels more conventional RD approaches based on local polynomial regression, which also involve bandwidth selection, but for a very different objective, namely finding an optimal balance between precision and bias at the threshold for local polynomials (e.g., [Ludwig and Miller, 2007](#); [Lee and Lemieux, 2010](#); [Imbens and Kalyanaraman, 2012](#)), whereas the objective in our framework is to find a subpopulation where our RD assumptions are plausible and the selected subpopulation defines the target population.

**3.4. Mode of inference.** Once the subpopulation  $\mathcal{U}_{s_0}$  is chosen, and under the RD assumptions 2-4, one can choose different modes of inference for the target causal estimands, as in the large literature of principal stratification. For example, under the additional Assumptions 5 and 6, the average causal effect for compliers,  $\tau_{s_0}$ , is non-parametrically point identified and could be estimated using standard moment-based (instrumental variable) methods. But the average causal effects for always-applicants and compliant-applicants,  $\tau_{AA,s_0}$  and  $\tau_{CA,s_0}$  can be only non-parametrically partially identified ([Mealli and Pacini, 2013](#)). One can also use likelihood approaches to parametrically estimate causal effects

(e.g., [Frumento et al., 2012](#); [Mercatanti, 2013](#)). Randomization-based inference ([Fisher, 1925](#)), as in [Cattaneo, Frandsen and Titiunik \(2015\)](#), could also be adopted.

In this article, we choose the Bayesian approach for inference for the following reasons. First, causal inference in RD designs usually involves complex observational data, with multiple sources of uncertainties, including the missing potential outcomes; the Bayesian approach is particularly useful for accounting for uncertainties and for pooling information from the data in such complex settings. Second, RD analysis usually relies on a sample of units with values of the forcing variable close to a single point, the size of which may be small; Bayesian methods, not relying on asymptotic approximations, are attractive in dealing with small samples. Third, in the Bayesian paradigm, the missing potential outcomes are treated as random variables, and all inferences are based on the posterior distributions of causal estimands, which are functions of potential outcomes. Thus inference about finite-sample and super-population estimands can be drawn using the same inferential procedures. Finally, pre-treatment variables can be easily incorporated in the Bayesian approach, which may improve efficiency of the analysis, i.e., reduce posterior variability.

**4. Bayesian inference.** Our development of the Bayesian approach builds on the seminal works of [Rubin \(1978\)](#) and [Imbens and Rubin \(1997\)](#). Below we give a brief outline for conduction principal stratification analysis using a Bayesian approach; the readers may refer to the existing literature for more details (e.g., [Elliott, Raghunathan and Li, 2010](#); [Schwartz, Li and Mealli, 2011](#); [Mattei, Li and Mealli, 2013](#)). Throughout the discussion, we use  $p(\cdot|\cdot)$  and  $\theta_{\cdot}$  to denote generic conditional distributions and the corresponding parameters, respectively.

Nine quantities are associated with each unit:  $Y_i(0)$ ,  $Y_i(1)$ ,  $W_i(0)$ ,  $W_i(1)$ ,  $A_i(0)$ ,  $A_i(1)$ ,  $\mathbf{X}_i$ ,  $Z_i$ ,  $S_i$ . Among these,  $S_i$  completely determines  $Z_i$ ; the principal stratum  $G_i = (A_i(0), A_i(1))$  and  $S_i$  completely determine  $(W_i(0), W_i(1))$ . Therefore, inference for causal effects involves only  $Y_i(0)$ ,  $Y_i(1)$ ,  $A_i(0)$ ,  $A_i(1)$ ,  $\mathbf{X}_i$ ,  $S_i$ , of which four are observed:  $S_i$ ,  $\mathbf{X}_i$ ,  $A_i^{obs} = A_i(Z_i)$ ,  $Y_i^{obs} = Y_i(Z_i)$ , and two are unobserved:  $A_i^{mis} = A_i(1 - Z_i)$ ,  $Y_i^{mis} = Y_i(1 - Z_i)$ .

Bayesian inference considers the observed values to be realizations of random variables and the unobserved values to be unobserved random variables. Let  $p(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{A}(0), \mathbf{A}(1), \mathbf{X}, \mathbf{S}; \mathcal{U}_{s_0})$  denote the joint probability density function of these random variables for all units in  $\mathcal{U}_{s_0}$ . We assume this distribution is unit-exchangeable, that is, it is invariant under a permutation of the unit indices. Then, with essentially no loss of generality, by appealing to de Finetti's theorem ([de Finetti, 1963](#)), we can assume that there exists an unknown parameter vector  $\theta$ , which is itself a random variable having a known prior distribution  $p(\theta)$  such that:

$$p(\mathbf{Y}(0), \mathbf{Y}(1), \mathbf{A}(0), \mathbf{A}(1), \mathbf{X}, \mathbf{S}; \mathcal{U}_{s_0}) = \int \prod_{i \in \mathcal{U}_{s_0}} p(Y_i(0), Y_i(1), A_i(0), A_i(1), \mathbf{X}_i, S_i | \theta) p(\theta) d\theta.$$

Bayesian inference of the causal estimands, which are functions of  $Y_i(z)$ 's and  $A_i(z)$ 's,

centers around deriving the posterior distribution for the parameter vector of their distribution, denoted by  $\theta_{Y,G}$ . Under Assumption 4, and assuming the parameters governing the distributions of the covariates, the forcing variable, and the potential outcomes are *a priori* distinct and independent, the posterior distribution of  $\theta_{Y,G}$  can be written as follows:

$$(6) \quad p(\theta_{Y,G} | \mathbf{Y}^{obs}, \mathbf{A}^{obs}, \mathbf{X}, \mathbf{S}; \mathcal{U}_{s_0}) \propto p(\theta_{Y|G}) \times p(\theta_G) \times \prod_{i \in \mathcal{U}_{s_0}} \left[ \int \int p(Y_i(0), Y_i(1) | G_i, \mathbf{X}_i; \theta_{Y|G}) p(G_i | \mathbf{X}_i; \theta_G) dY_i^{mis} dA_i^{mis} \right].$$

The above decomposition suggests that two models need to be specified for model-based inference: (1) the model for potential outcomes conditional on principal strata and covariates, and (2) the model for principal strata conditional on covariates, as well as the prior distribution for the parameters,  $p(\theta_{Y,G})$ , with  $\theta_{Y,G} = (\theta_G, \theta_{Y|G})$ .

Let  $\pi_{i,g} = \Pr(G_i = g | \mathbf{X}_i; \theta_G)$  and  $f_{i,gz} = p(Y_i(z) | G_i = g, \mathbf{X}_i; \theta_{Y|G})$ . Then the posterior distribution of  $\theta_{Y,G}$  given the observed data can be written as follows:

$$(7) p(\theta_{Y,G} | \mathbf{Y}^{obs}, \mathbf{A}^{obs}, \mathbf{X}, \mathbf{S}; \mathcal{U}_{s_0}) \propto p(\theta_{Y,G}) \times \prod_{i \in \mathcal{U}_{s_0}: S_i > s_0, A_i^{obs} = 0} (\pi_{i,CA} f_{i,CA,0} + \pi_{i,NA} f_{i,NA}) \times \prod_{i \in \mathcal{U}_{s_0}: S_i > s_0, A_i^{obs} = 1} \pi_{i,AA} f_{i,AA,0} \\ \times \prod_{i \in \mathcal{U}_{s_0}: S_i \leq s_0, A_i^{obs} = 0} \pi_{i,NA} f_{i,NA} \times \prod_{i \in \mathcal{U}_{s_0}: S_i \leq s_0, A_i^{obs} = 1} (\pi_{i,AA} f_{i,AA,1} + \pi_{i,CA} f_{i,CA,1}),$$

where  $f_{i,NA} = f_{i,NA,0} = f_{i,NA,1}$  by the exclusion restriction (Assumption 6). The likelihood function, specified by the four products, does not depend on the association between the potential outcomes  $Y_i(0)$  and  $Y_i(1)$ . Therefore the posterior distribution of the association parameters equal their prior distribution as long as the association parameters are *a priori* independent of the other parameters, as we assume henceforth. The population-average causal estimands  $\tau_{AA,s_0}$ ,  $\tau_{CA,s_0}$ , and  $\tau_{s_0}$  are functions of the parameter vector  $\theta_{Y,G}$ , which is free of the association parameters, therefore inference for them does not involve the association parameters (also see discussion in Imbens and Rubin, 1997). Inference for sample-average causal estimands for the units in the study, on the other hand, do generally involve the association parameters. In our application inference for sample-average causal estimands is drawn under the assumption that for each unit  $i$ , potential outcomes,  $Y_i(0)$  and  $Y_i(1)$ , are independent conditional on  $\mathbf{X}_i$  and  $\theta$ .

Posterior inference of  $\theta_{Y,G}$  can be obtained using Gibbs sampling with a data augmentation step to impute the missing  $A_i^{mis}$ , iteratively drawing from the two posterior predictive distributions,  $p(\theta_{Y,G} | \mathbf{Y}^{obs}, \mathbf{A}^{obs}, \mathbf{A}^{mis}, \mathbf{X}, \mathbf{S}; \mathcal{U}_{s_0})$  and  $p(\mathbf{A}^{mis} | \mathbf{Y}^{obs}, \mathbf{A}^{obs}, \mathbf{X}, \mathbf{S}, \theta_{Y,G}; \mathcal{U}_{s_0})$ .

Specification of  $\pi_{i,g}$ ,  $f_{i,gz}$  and corresponding prior to posterior computation depends on the specific application. Details of the models and computation in our application will be

provided in Section 5. As a general guideline, we recommend to specify  $\pi_{i,g}$  and  $f_{i,gz}$  conditional on both covariates  $\mathbf{X}$  and the forcing variable  $S$ , even though Equation (6) suggests conditioning on  $S$  is not required. Indeed, if the true subpopulations  $\mathcal{U}_{s_0}$  were known, in theory, we would not need to adjust for  $S$ , because local randomization guarantees that for units in  $\mathcal{U}_{s_0}$  values of the forcing variable falling above or below the threshold are independent of the potential outcomes. However, in practice, the true subpopulations  $\mathcal{U}_{s_0}$  are usually unknown and once a subpopulation has been selected, that is, once a value for  $h$ , say  $h^*$ , has been chosen, there may be some units with a realized value of  $S$  between  $s_0 - h^*$  and  $s_0 + h^*$  who do not belong to  $\mathcal{U}_{s_0}$ . For these units there may be a relationship between the forcing variable and potential outcomes, and these potential dependences need to be modeled. Specifically, systematic differences in the forcing variable  $S$  that, by definition, occur between eligible and ineligible units, may affect inference in the presence of students who do not belong to  $\mathcal{U}_{s_0}$ .

## 5. Evaluation of Italian university grants.

**5.1. Data.** We apply the proposed method to the data from the cohort of first-year students enrolled in 2004 to 2006 at University of Pisa and University of Florence. For each student, information on grant application status ( $A_i^{obs}$ ), grant receipt status ( $W_i^{obs}$ ) at the beginning of the academic year, dropout status at the end of the academic year, and covariates ( $\mathbf{X}_i$ ) is obtained from ministry of education and university administrative records. The forcing variable  $S$  is a combined economic measure of each student, calculated from one's income tax return and property adjusted for family size based on a formula that is typically not fully known to the students. In all three years, the threshold of eligibility is the combined economic measure of a student below 15 000 euros. Thus, the eligibility status ( $Z$ ) is also observed. Typically, students need support from fiscal experts to compute their value of  $S$ , and the income revenue authority conducts random inspections to verify that the official tax return were reported. These factors make extremely difficult, if not impossible, for students or students' families to manipulate the value of  $S$  in order to end up on the right side of the threshold. Therefore we argue that the local randomization assumption is reasonable here. Ineligible students apply either usually because they are not fully aware of their eligibility status, or because they hope that their application will be still considered because of extra funding or other considerations.

Covariates include sex, high school grade, high school type (4 categories), major in university (6 categories), indicator of year of enrollment (2004, 2005, 2006) and indicator of university (Pisa vs. Florence). Note that the data only include students who had a high school grade of at least 70/100 and applied either for a grant or for a reduction of tuition fee. Summary statistics of important variables for the students with the combined economic measure  $S$  within 1000 euros of the threshold are given in Table 1. An unadjusted comparison would suggest that the applicants have higher high-school grades, which is an important indicator of a student's academic performance, but also higher dropout rate



regardless of their eligibility status.

Application rate and dropout rate as a function of  $S$  among the students are given in Figure 1. The overall dropout rate is high, consistently between 30% to 50% regardless of the economic measure. From the fitted lines using local logistic polynomial models with order 3 on the two sides of the threshold, discontinuity is clearly visible in both application rate and dropout rate at the threshold. As the economic measure increases, application rate steadily decreases, while the trend in dropout rate has a concave change at the threshold, increasing on the left of the threshold and decreasing on the right.

TABLE 1

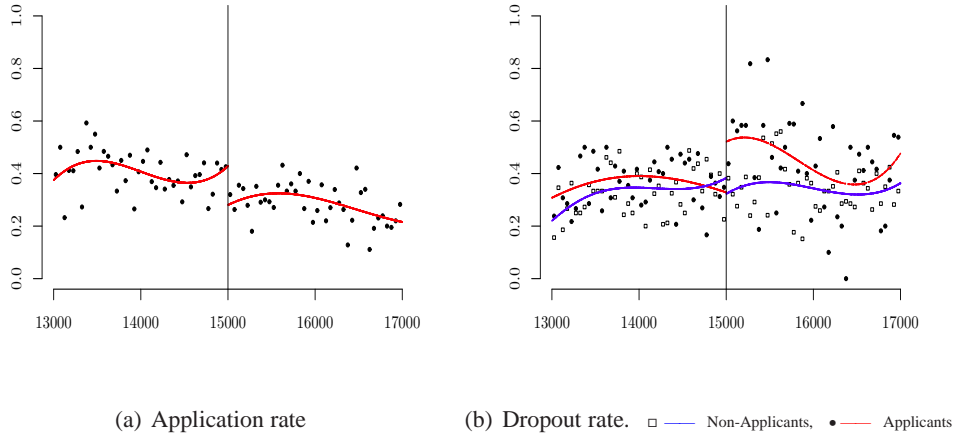
Summary statistics of the first-year students enrolled in 2004 – 2006 at Universities of Pisa and Florence, for the students with  $S_i \in (14\,000, 16\,000)$  euros (i.e.,  $h = 1\,000$ ,  $s_0 = 15\,000$ ).

Variable	Z = 0		Z = 1	
	$A^{obs} = 0$	$A^{obs} = 1$	$A^{obs} = 0$	$A^{obs} = 1$
Sample Size	657	304	703	444
Dropout	0.36	0.50	0.35	0.36
$S$ (euros)	15 495	15 509	14 504	14 499
Female	0.59	0.61	0.60	0.55
HS Grade	80.80	84.35	80.17	84.47
University (Pisa)	0.37	0.51	0.37	0.51

**5.2. Selection of the subpopulation.** We apply the Bayesian approach to multiple testing discussed in Section 3.3 to find subpopulations of units where our RD assumptions hold. Specifically we use a hierarchical Bayesian model for assessing the balance of the covariates between eligibility groups. We specify probit models for binary variables; conditional probit models for categorical variables and Gaussian models for continuous variables. Formally, we assume that  $X_j \sim N(\gamma_{0j} + \gamma_{1j}Z_i, \sigma_j^2)$  if  $X_j$  is continuous, and  $\Pr(X_{ij} = 1) = \Pr(X_{ij}^* > 0)$  with  $X_{ij}^* \sim N(\gamma_{0j} + \gamma_{1j}Z_i, 1)$ , if  $X_j$  is binary. If  $X_j$  is a categorical variable taking on  $K$  values we assume that  $\Pr(X_{ij} = 1) = \Pr(X_{ij}^{*(1)} \leq 0)$ , and  $\Pr(X_{ij} = k) = \Pr(\bigcap_{\ell=1}^{k-1} \{X_{ij}^{*(\ell)} > 0\} \cap X_{ij}^{*(k)} \leq 0)$  for  $k = 2, \dots, K-1$ , where  $X_{ij}^{*(k)} \sim N(\gamma_{0j}^{(k)} + \gamma_{1j}^{(k)}Z_i, 1)$ ,  $k = 1, \dots, K-1$ , independently. Let  $\gamma_{0j} = (\gamma_{0j}^{(1)}, \dots, \gamma_{0j}^{(K-1)})'$  and  $\gamma_{1j} = (\gamma_{1j}^{(1)}, \dots, \gamma_{1j}^{(K-1)})'$ .

We specify the following prior distributions for the model parameters. The variances of the continuous variables have an inverse-Gamma distribution:  $\sigma_j^2 \sim IG(\underline{a}, \underline{b})$ . The  $\gamma_0$ 's have Gaussian prior distributions: for continuous and binary variables,  $\gamma_{0j} \sim N(\mu_{\gamma_0}, \sigma_{\gamma_0}^2)$ , and for categorical variables,  $\gamma_{0j} \sim N(\mu_{\gamma_0} \mathbf{u}_{K-1}, \sigma_{\gamma_0}^2 \mathbf{I}_{K-1})$  with  $\mathbf{u}_{K-1}$  and  $\mathbf{I}_{K-1}$  being the  $K-1$ -dimensional vector of ones and the identity matrix of order  $K-1$ , respectively. Further, for continuous and binary variables, parameters  $\gamma_{1j}$  are the difference between means/proportions for eligible and ineligible units. If  $\gamma_{1j} = 0$  then  $X_j$  has the same distribution for eligible and ineligible units. For a categorical variable taking on  $K$  values, the proportion of units in each category is the same for eligible and ineligible units if and only

FIG 1. Application rate (a) and dropout rate (b) as a function of the forcing variable for the first year students in Universities of Florence and Pisa in 2004–2006. The smoothed lines are estimated using polynomial logistic regression models (of order 3) on each side of the threshold separately; each point are calculated from the units within a binwidth of 50 euros.



if  $\gamma_{1j}^{(k)} = 0$  for each  $k = 1, \dots, K - 1$ . We assign positive probability to these possibilities using the following mixture prior distributions:

$$\gamma_{1j} \sim \pi_{\gamma_1} \delta_0(\gamma_{1j}) + (1 - \pi_{\gamma_1}) N(\mu_{\gamma_1}, \sigma_{\gamma_1}^2)$$

and

$$\gamma_{1j} \sim \prod_{k=1}^{K-1} \left[ \pi_{\gamma_1} \delta_0(\gamma_{1j}^{(k)}) + (1 - \pi_{\gamma_1}) N(\mu_{\gamma_1}, \sigma_{\gamma_1}^2) \right],$$

where  $\delta_0(\cdot)$  is the Dirac delta distribution.

For the hyperparameters, we assign the following prior distributions:  $\mu_{\gamma_0} \sim N(\underline{\mu}_{\gamma_0}, \underline{\sigma}_{\gamma_0}^2)$ ;  $\sigma_{\gamma_0}^2 \sim IG(\underline{a}_{\gamma_0}, \underline{b}_{\gamma_0})$ ;  $\mu_{\gamma_1} \sim N(\underline{\mu}_{\gamma_1}, \underline{\sigma}_{\gamma_1}^2)$ ;  $\sigma_{\gamma_1}^2 \sim IG(\underline{a}_{\gamma_1}, \underline{b}_{\gamma_1})$ ; and  $\pi_{\gamma_1} \sim Beta(\underline{a}_{\pi}, \underline{b}_{\pi})$ .

We implement the Bayesian model for assessing the balance of covariates on the two sides of the threshold for various subpopulations defined by different values of  $h$ . Details on the Monte Carlo Markov Chain (MCMC) for the posterior computation are relegated to

[Web Supplementary Material](#). Table 2 shows the posterior probabilities that the covariates have the same distribution between eligible and ineligible students for the subpopulations defined by  $h = 250, 500, 750, 1\,000, 1\,500, 2\,000, 2\,500, 3\,000, 4\,000, 5\,000$ . These values show that the probability of the pre-assignment variables being well balanced is high for subpopulations defined by values of  $h$  strictly lower than 1 500: the vast majority of these probabilities are larger than or close to 0.8. Note that the probabilities are in general lower among the covariates of “major in university”, suggesting these covariates may not be as balanced as other covariates. Nonetheless, nearly all these probabilities are still higher 0.6 with a single lowest probability being 0.565 (Tech major in university). For larger subpopulations some covariates, such as the “indicator of university,” are clearly unbalanced.

Given that the risk that a chosen interval around the threshold defines a subpopulation that includes units not belonging to the target subpopulation,  $\mathcal{U}_{s_0}$ , is not zero, in order to account for the presence of these units, we conduct the subsequent analyses conditioning on both covariates and the realized values on the forcing variable. Also we evaluate the robustness of our results conducting analyses using various values of  $h$  ( $h = 500, 1\,000, 1\,500$ )

**5.3. Parametric models.** For the units within the selected subpopulation  $\mathcal{U}_{s_0}$ , we assume parametric models for the outcome ( $f_{gz}$ ) and principal strata ( $\pi_g$ ). Alternative models, such as Student- $t$  models ([Chib and Jacobi, 2011](#)) and Bayesian nonparametric models ([Schwartz, Li and Mealli, 2011](#)), can be considered. Note that although we are using parametric models, identification does not rely on parametric assumptions. The model for the principal strata of application consists of two conditional probit models:

$$\begin{aligned}\pi_{i,AA} &= \Pr(G_i^*(AA) \leq 0), \\ \pi_{i,NA} &= \Pr(G_i^*(AA) > 0 \text{ and } G_i^*(NA) \leq 0), \\ \pi_{i,CA} &= 1 - \pi_{i,AA} - \pi_{i,NA},\end{aligned}$$

where

$$G_i^*(AA) = \alpha_{AA,0} + \alpha_{AA}^{(S)} S_i^* + \mathbf{X}_i' \alpha_{AA}^{(X)} + \epsilon_{AA,i}, \quad G_i^*(NA) = \alpha_{NA,0} + \alpha_{NA}^{(S)} S_i^* + \mathbf{X}_i' \alpha_{NA}^{(X)} + \epsilon_{NA,i},$$

with  $\epsilon_{AA,i} \sim N(0, 1)$ ,  $\epsilon_{NA,i} \sim N(0, 1)$  independently, and  $S_i^* = (S_i - s_0)/1000$ .

Dropout, the primary outcome in our application, is binary. Therefore, we assume the following generalized linear outcome model with a probit link ([Albert and Chib, 1993](#)):

$$\Pr(Y_i(z) = 1 | G_i = g, S_i, \mathbf{X}_i) = \Phi(\beta_{0,g,z} + \beta_{g,z}^{(S)} S_i^* + \mathbf{X}_i' \beta_{g,z}^{(X)}).$$

We impose prior equality of the slope coefficients in the outcome regressions:  $\beta_{g,z}^{(X)} \equiv \beta^{(X)}$  for  $g = AA, CA, NA$  and  $z = 0, 1$ .

Define  $\alpha_g = [\alpha_{g0}, \alpha_g^{(S)}, \alpha_g^{(X)}]'$ ,  $g = AA, NA$ , and  $\beta_{g,z} = [\beta_{0,g,z}, \beta_{g,z}^{(S)}]'$ ,  $g = AA, CA, NA$ ;  $z = 0, 1$ . By Assumption 6,  $\beta_{NA,0} = \beta_{NA,1}$ . We assume that parameters are *a priori* independent and use multivariate normal prior distributions:

$$\alpha_g \sim N(\underline{\mu}_{\alpha_g}; \underline{\sigma}_{\alpha_g}^2 \mathbf{I}), \quad \beta_{g,z} \sim N(\underline{\mu}_{\beta_{g,z}}; \underline{\sigma}_{\beta_{g,z}}^2 \mathbf{I}), \quad \beta^{(X)} \sim N(\underline{\mu}_{\beta}; \underline{\sigma}_{\beta}^2 \mathbf{I})$$

where  $\mathbf{I}$  is the identity matrix. We specify flat priors setting the hyper-parameters as follows: setting  $\underline{\mu}_{\alpha_g}, \underline{\mu}_{\beta_{g,z}}, \underline{\mu}_{\beta}$  to be null vectors; and setting large prior variances  $\underline{\sigma}_{\alpha_g}^2 = 10$ ,  $\underline{\sigma}_{\beta_{g,z}}^2 = 10$ ,  $\underline{\sigma}_{\beta}^2 = 10$  for  $g = AA, CA, NA$ ;  $z = 0, 1$ .

**5.4. Posterior computation.** Details of the MCMC algorithm for the posterior computation based on the outline in Section 4 are given in [Web Supplementary Material](#). Upon obtaining the posterior draws of the parameters, we calculate three estimates for each causal estimand: population-average effect *within*  $\mathcal{U}_{s_0}$  and *at*  $s_0$ , and sample-average effect within  $\mathcal{U}_{s_0}$ . The population-average effects within  $\mathcal{U}_{s_0}$  are calculated averaging the model-based dropout proportions over the empirical distribution of the pre-assignment variables and the forcing variable:

$$\frac{\sum_{i \in \mathcal{U}_{s_0}} \pi_{i,g} \Phi(\beta_{0,g,1} + \beta_{g,1}^{(S)} S_i^* + \mathbf{X}_i' \beta^{(X)})}{\sum_{i \in \mathcal{U}_{s_0}} \pi_{i,g}} - \frac{\sum_{i \in \mathcal{U}_{s_0}} \pi_{i,g} \Phi(\beta_{0,g,0} + \beta_{g,0}^{(S)} S_i^* + \mathbf{X}_i' \beta^{(X)})}{\sum_{i \in \mathcal{U}_{s_0}} \pi_{i,g}},$$

for  $g = AA, CA, \{AA, CA\}$ . The population-average effects at  $s_0$  are calculated in a similar way setting  $S_i^* = 0$  (i.e.,  $S_i = s_0$ ) for each  $i$ . To obtain the sample-average estimates, we compute the posterior predictive distributions of the potential outcomes for each student  $i$  in  $\mathcal{U}_{s_0}$ , based on which the sample average is calculated.

**5.5. Results.** We conducted Bayesian analysis using  $h = 500, 1\,000, 1\,500$ . Posterior inference is based on 5 000 draws from the posterior distributions simulated using single chains, which were run for 125 000 iterations. To assess convergence of iterative simulation methods, we calculated the Cramer-von-Mises statistic to test the null hypothesis that the sampled values come from a stationary distribution and visual inspected the trace-plots of the causal parameters (functions of model parameters). We also run multiple MCMC chains with different starting for each  $h$  to evaluate the mixing of the chains using the Gelman-Rubin statistic ([Gelman and Rubin, 1992](#)). The results provided no evidence against convergence<sup>1</sup>.

<sup>1</sup>We also conducted Bayesian analysis using alternative models with different order polynomials in  $S$  as well as models conditioning only on  $S$  (without using the pretreatment variables) and null models, conditioning on neither  $S$  nor the pre-treatment covariates. Consistently to results found in [Mealli and Rampichini \(2012\)](#), higher order polynomials do not lead to substantial inferential benefits, and posterior distributions of the causal effects of interest did not substantially change with the alternative models, so here we only show the results based on models conditioning on both  $S$  and the pre-treatment covariates.

Table 3 shows posterior medians and 95% credible intervals for the principal strata proportions under monotonicity and for the causal parameters  $\tau_{AA,s_0}$ ,  $\tau_{CA,s_0}$ ,  $\tau_{s_0}$ , for bandwidths ranging from 500 to 1500 euros. The results are robust across different bandwidths. The estimated proportions of the principal strata are very similar across different  $h$ : there are more than 61% never-applicants, more than 32% always-applicants and less than 6.5% compliant-applicants. The three estimates for the same causal parameter are also similar. The posterior distributions of the causal effect for always-applicants,  $\tau_{AA,s_0}$ , and the union of always-applicants and compliant-applicants,  $\tau_{s_0}$ , are centered on negative values, and the 95% credible intervals do not cover 0, irrespective of the choice of the bandwidth.

For instance, consider the finite-sample causal effects for the subpopulation within  $h = 1000$  euros around the threshold (middle block of columns in Table 3). The estimated  $\tau_{s_0}$  suggests a 13.9% (95% CI: (3.4%; 24.7%)) reduction in dropout rate for the students who receive the grants. The estimated  $\tau_{AA,s_0}$  suggests an even stronger positive effect among the always-applicants: a 16.1% (95% CI: (5%; 27%)) reduction in dropout rate. In fact,  $\tau_{s_0}$ , which is a weighted average of the effects for always-applicants and compliant-applicants, appears to be diluted by the somewhat surprising small effect among the compliant-applicant. However, the data do not seem to contain much information on compliant-applicants (the estimated proportion of compliant-applicants is very small, less than 5%), and the effects were estimated with large uncertainties.

These results suggest that the current Italian university grants are effective in reducing dropout from universities among students from families with annual economic measure around 15 000 euros. Our analysis also reveals some additional information for policy making. Specifically, always-applicants and compliant-applicants are found to be heterogeneous with respect to the effect of the grants. The causal effect for compliers,  $\tau_{s_0}$ , usually estimated in a standard IV analysis that ignores the application information, is attenuated by the small (and negative) effect estimated for the small proportion of compliant-applicants. From a cost-effective perspective, it appears more beneficial for education administrations to lower the eligibility criteria (i.e., decrease the threshold  $s_0$ ) to allow more applicants to get the grant, than to increase the amount of the grant to awardees. The combination of low percentage of compliant-applicants and high percentage of always-applicants suggests that most students with the economic measure being around the threshold who intend to apply for the grants would apply irrespective of their eligibility. From a policy perspective, this implies that educational administrations should better explain the rule of eligibility to potential applicants to discourage ineligible students from applying, and thus reduce unnecessary efforts from these students and the administration, for processing these applications.

*5.6. Posterior Predictive Model Checking.* Assessing the plausibility of model assumptions is critical in model-based approaches. Model checking here is not as crucial as in other model-based approaches thanks to the randomization assumption, but it is

still prudent to check the model fit since there are uncertainties in the selection of  $\mathcal{U}_{s_0}$ . We adopt Bayesian posterior predictive checks (Gelman, Meng and Stern, 1996) to assess goodness-of-fit of our models in the application. Posterior predictive checks evaluate goodness-of-fit of models by measuring the discrepancy between the observed data and replicated data simulated from its posterior predictive distribution. The particular procedure adopted here is similar to that in Mattei, Li and Mealli (2013, Section 6). Specifically, we consider three discrepancy measures aim at assessing whether the model can preserve broad features of signal, noise and signal-to-noise ratio (SNR) in the drop-out status distribution for compliant-applicants, always-applicants and the union of these two principal strata, and calculate posterior predictive  $p$ -values (PPPVs) to summarize discrepancies between the observed data and replicated data. Extreme (close to 0 or 1) PPPVs can be interpreted as evidence of lack-of-fit of the model in, at least some aspects of, the observed data. Further details of the procedure are relegated in [Web Supplementary Material](#).

Table 4 shows the PPPVs for the model-fit to the subpopulation with bandwidth of 500, 1 000 and 1 500 euros, respectively. The PPPVs suggest good model-fit for all bandwidths, except for a slight under-fit for always-applicants in the subpopulation with  $h = 500$ , which is possibly due to the small sample size. We have also calculated the less conservative sampled posterior predictive  $p$ -values (Johnson, 2007; Gosselin, 2011) and obtained similar conclusions.

**6. Discussion.** Motivated from the evaluation of Italian university grants, we propose a probabilistic formulation of the assignment mechanism for regression discontinuity designs and develop a full Bayesian approach to draw causal inference within the framework of principal stratification. In particular, we illustrate how to utilize information on application status to gain additional insights in program evaluation. Applying the method to the data from two Italian universities, we find university grants reduce dropping out of higher education for students from low-income families and the effect size is especially pronounced for motivated students (always-applicants).

The core of the approach we propose is the local randomization assumption (Assumption 4), which is intrinsically non-testable. Therefore, it may be worthwhile to conduct sensitivity analyses aimed at assessing the robustness of the results with respect to violations of the local randomization assumption. To this end, we conduct further analyses deriving the posterior distributions of the causal estimands of interest under three additional model specifications: (1) a model where we specify the model for principal strata,  $\pi_{i,g}$ , and the conditional model for potential outcomes given principal strata,  $f_{i,gz}$ , conditioning on neither the forcing variable nor the pre-treatment variables; (2) a model where we specify  $\pi_{i,g}$  and  $f_{i,gz}$  conditioning only on the forcing variable, without including the pre-treatment variables; and (3) a model where we specify  $\pi_{i,g}$  and  $f_{i,gz}$  conditioning only on the pre-treatment variables, without including the forcing variable. Under local randomization, adjusting inference for either the forcing variable,  $S$ , or the pre-treatment



variables,  $\mathbf{X}$ , should not be required, therefore we expect that results are similar across different model specifications. Indeed results, shown in [Web Supplementary Material](#), are robust across different model specifications, suggesting that causal inference under the local randomization assumption is credible and fully defensible.

A fundamental distinction between our approach and the previous local-regression based RD approaches lies in the role of the forcing variable in the analysis. Specifically, previous approaches generally view the forcing variable as a pre-assignment covariate rather than a random variable as in our approach. As a consequence, the standard overlap assumption, which requires that there are both treated and control units for all values of the covariates including the forcing variable, is violated. Violation of the overlap assumption implies that the conditional independence assumption, which trivially holds in RD settings, cannot be exploited directly. Instead some kind of extrapolation is required, and in order to avoid that estimates heavily rely on extrapolation, previous analyses focus on causal effects of the treatment for units *at* the threshold. Smoothness assumptions, for example, continuity of conditional regression functions of potential outcomes given the forcing variable, are usually assumed to draw inference on those causal effects. Local randomization and continuity are different assumptions, leading to different causal estimands: under continuity assumptions units with a realized value of the forcing variable around the threshold are used to draw inference on causal effects for units *at* the threshold, whereas under local randomization, inference is drawn on causal effects for units *around* the threshold.

In the evaluation of Italian university grants, other than dropout, student's academic performance (measured by total credits taken or passing rate of exams) is also of great interest in policy. As illustrated by [Mattei, Li and Mealli \(2013\)](#) and [Mercatanti, Li and Mealli \(2014\)](#), jointly modeling two outcomes, dropout and academic performance in this case, would be worthwhile for both practical and inferential purposes, and it is at the top of our research agenda.

After the first year, the Italian university grant assignment rule combines sequential and RD designs ([Cellini, Ferreira and Rothstein, 2010](#)): grants are allocated both on the basis of students family economic indicator and on the ground of their academic performance (exam scores above a certain threshold). Such complex assignment mechanisms pose challenges to causal inference, requiring new structures and assumptions; meanwhile, they also present great opportunities for extending the existing framework to more general RD settings. One specific direction of our future research is to develop methods that combine Bayesian tools for RDs and dynamic treatment regimes ([Murphy, 2003](#); [Zajonc, 2012](#)) in the presence of multiple forcing variables ([Imbens and Zajonc, 2011](#)).

**Acknowledgements.** The authors are grateful to the associate editor and two reviewers for constructive comments, to Quanli Wang for computing support, to Guido Imbens, Michael Hudgens and Sid Chib for helpful discussions.

## SUPPLEMENTARY MATERIAL

**Web Supplementary Material: Details of Caculation and Sensitivity Analysis**

([http://lib.stat.cmu.edu/aoas/???/???/](http://lib.stat.cmu.edu/aoas/???/???)). We describe in detail the Bayesian approach we used to select the subpopulations, the Markov Chain Monte Carlo (MCMC) methods used to simulate the posterior distributions of the parameters of the models, the posterior predictive checks, and the sensitivity analysis regarding local randomization described in Section 6.

**References.**

- ALBERT, J. and CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association* **88** 669-679.
- ANGRIST, J. D., IMBENS, G. W. and RUBIN, D. B. (1996). Identification of causal effects using instrumental variables (with discussion). *Journal of the American Statistical Association* **91** 444-472.
- BARNARD, J., FRANGAKIS, C. F., HILL, J. L. and RUBIN, D. B. (2003). Principal stratification approach to broken randomized experiments: A case study of school choice vouchers in New York City (with discussion). *Journal of the American Statistical Association* **98** 299-323.
- BATTISTIN, E. and RETTORE, E. (2008). Ineligibles and eligible non-participants as a double comparison group in regression discontinuity designs. *Journal of Econometrics* **142** 715-730.
- BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society - Series B* **57** 289-300.
- BERK, R. A. and DE LEUW, J. (1999). An evaluation of California's inmate classification system using a generalized regression discontinuity design. *Journal of the American Statistical Association* **94** 1045-1052.
- BERRY, S. M. and BERRY, D. A. (2004). Accounting for Multiplicities in Assessing Drug Safety: A Three-Level Hierarchical Mixture Model. *Biometrics* **60** 418-426.
- CATTANEO, M. D., FRANDSEN, B. and TITIUNIK, R. (2015). Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. *Journal of Causal Inference* **3**(1) 1-24.
- CELLINI, S. R., FERREIRA, F. and ROTHSTEIN, J. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. *Quarterly Journal of Economics* **125** 215-261.
- CHIB, S. and GREENBERG, E. (2014). Nonparametric Bayes Analysis of the Sharp and Fuzzy Regression Discontinuity Designs Technical Report, Washington Univesity at St Louis, Olin School of Business.
- CHIB, S. and JACOBI, L. (2011). Returns to Compulsory Schooling in Britain: Evidence from a Bayesian Fuzzy Regression Discontinuity Analysis. IZA Discussion Paper No. 5564.
- COOK, T. D. (2008). Waiting for life to arrive: A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics* **142** 636-654.
- DE FINETTI, B. (1963). Foresight: its logical laws, its subjective sources. In *Studies in Subjective Probability*.
- DINARDO, J. and LEE, D. S. (2011). *Handbook of Labor Economics* **4A** Program Evaluation and Research Designs 463-536. Elsevier Science B.V.
- ELLIOTT, M., RAGHUNATHAN, T. and LI, Y. (2010). Bayesian inference for causal mediation effects using principal stratification with dichotomous mediators and outcomes. *Biostatistics* **11** 353-372.
- FISHER, R. A. (1925). *Statistical Methods for Research Workers. First edition*. Oliver and Boyd, Edimburgh.
- FRANGAKIS, C. E. and RUBIN, D. B. (2002). Principal stratification in causal inference. *Biometrics* **58** 191-199.
- FRUMENTO, P., MEALLI, F., PACINI, B. and RUBIN, D. (2012). Evaluating the Effect of Training on Wages in the Presence of Noncompliance, Nonemployment, and Missing Outcome Data. *Journal of the American Statistical Association* **107** 450-466.
- GARIBALDI, P., GIAVAZZI, F., ICHINO, A. and RETTORE, E. (2012). College Cost and Time to Complete a Degree: Evidence from Tuition Discontinuities. *The Review of Economics and Statistics* **94** 699-711.
- GELMAN, A. E., MENG, X. L. and STERN, H. S. (1996). Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica* **6** 733-807.

- GELMAN, A. and RUBIN, D. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **7** 457-472.
- GOSSELIN, F. (2011). A New Calibrated Bayesian Internal Goodness-of-Fit Method: Sampled Posterior p-Values as Simple and General p-Values That Allow Double Use of the Data. *PLoS ONE* **6** 1-10.
- HAHN, J., TODD, P. E. and VAN DER KLAUW, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* **69** 201-209.
- IMBENS, G. W. (2004). Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics* **86** 4-29.
- IMBENS, G. W. and ANGRIST, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica* **62** 467-476.
- IMBENS, G. W. and KALYANARAMAN, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies* **79** 933-959.
- IMBENS, G. W. and LEMIEUX, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* **142** 615-635.
- IMBENS, G. W. and RUBIN, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *The Annals of Statistics* **25** 305-327.
- IMBENS, G. W. and RUBIN, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.
- IMBENS, G. W. and ZAJONC, T. (2011). Regression discontinuity design with multiple forcing variables Technical Report, Harvard University, Department of Economics.
- JOHNSON, V. E. (2007). Bayesian model assessment using pivotal quantities. *Bayesian Analysis* **2** 719-734.
- LEE, D. S. (2008). Randomized Experiments from Non-random Selection in U.S. House Elections. *Journal of Econometrics* **142** 675-697.
- LEE, D. S. and LEMIEUX, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* **48** 281-355.
- LUDWIG, J. and MILLER, D. L. (2007). Does head start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics* **122** 15981-208.
- MATTEI, A., LI, F. and MEALLI, F. (2013). Exploiting multiple outcomes in Bayesian principal stratification analysis with application to the evaluation of a job training program. *The Annals of Applied Statistics* **7** 2360-2013.
- MEALLI, F. and PACINI, B. (2013). Using secondary outcomes to sharpen inference in randomized experiments with noncompliance. *Journal of the American Statistical Association* **108** 1120-1131.
- MEALLI, F. and RAMPICHINI, C. (2012). Evaluating the effects of university grants using regression discontinuity designs. *Journal of the Royal Statistical Society, Series A* **175** 775-798.
- MEALLI, F. and RUBIN, B. D. (2002). Assumptions when Analyzing Randomized Experiments with Noncompliance and Missing Outcomes. *Health Services & Outcomes Research Methodology* **3** 225-232.
- MERCATANTI, A. (2013). A Likelihood-Based Analysis for Relaxing the Exclusion Restriction in Randomized Experiments With Noncompliance. *Australian and New Zealand Journal of Statistics* **55** 129-153.
- MERCATANTI, A., LI, F. and MEALLI, F. (2014). Evaluating the effects of university grants using regression discontinuity designs. *Statistical Analysis and Data Mining* **8**(1) 34-48.
- MURPHY, S. A. (2003). Optimal Dynamic Treatment Regimes. *Journal of the Royal Statistical Society, Series B* **65** 331-366.
- RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66** 688-701.
- RUBIN, D. B. (1978). Bayesian inference for causal effects. *The Annals of Statistics* **6** 34-58.
- RUBIN, D. B. (1980). Discussion of "Randomization analysis of experimental data: the Fisher randomization test" by D. Basu. *Journal of the American Statistical Association* **75** 591-593.
- SALES, A. and HANSEN, B. (2014). Limitless Regression Discontinuity: Causal Inference for a Population Surrounding a Threshold. *arXiv:1403.5478*.
- SCHWARTZ, S. L., LI, F. and MEALLI, F. (2011). A Bayesian Semiparametric Approach to Intermediate Variables

- in Causal Inference. *Journal of the American Statistical Association* **31** 949-962.
- SCOTT, J. G. and BERGER, J. O. (2006). An exploration of aspects of Bayesian multiple testing. *Journal of Statistical Planning and Inference* **136** 2144-2162.
- THISTLETHWAITE, D. and CAMPBELL, D. (1960). Regression-discontinuity analysis: an alternative to the ex-post facto experiment. *Journal of Educational Psychology* **51** 309-317.
- VAN DER KLAUW, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review* **43** 1249-1287.
- VAN DER KLAUW, W. (2008). Regression-discontinuity analysis: A survey of recent development in economics. *Labour* **22** 219-245.
- ZAJONC, T. (2012). Bayesian Inference for Dynamic Treatment Regimes: Mobility, Equity, and Efficiency in Student Tracking. *Journal of the American Statistical Association* **107** 80-92.

ALESSANDRA MATTEI AND FABRIZIA MEALLI  
UNIVERSITY OF FLORENCE  
DEPARTMENT OF STATISTICS, INFORMATICS, APPLICATIONS  
VIALE MORGAGNI, 59  
50134 FLORENCE, ITALY  
E-MAIL: [mattei@disia.unifi.it](mailto:mattei@disia.unifi.it)  
[mealli@disia.unifi.it](mailto:mealli@disia.unifi.it)

FAN LI  
DUKE UNIVERSITY  
DEPARTMENT OF STATISTICAL SCIENCE  
122 OLD CHEMISTRY BUILDING  
DURHAM, NORTH CAROLINA 27708-0251, U.S.A.  
E-MAIL: [fli@stat.duke.edu](mailto:fli@stat.duke.edu)

TABLE 2. *Posterior probabilities that the covariates have the same distribution between eligible and ineligible students for various subpopulation*

Variable	<i>h</i> =250 ( <i>n</i> =528)	<i>h</i> =500 ( <i>n</i> =1 042)	<i>h</i> =750 ( <i>n</i> =1 577)	<i>h</i> =1 000 ( <i>n</i> =2 108)	<i>h</i> =1 500 ( <i>n</i> =3 166)	<i>h</i> =2 000 ( <i>n</i> =4 197)	<i>h</i> =2 500 ( <i>n</i> =5 159)	<i>h</i> =3 000 ( <i>n</i> =6 113)	<i>h</i> =4 000 ( <i>n</i> =8 061)	<i>h</i> =5 000 ( <i>n</i> =9 846)
Sex	.955	.950	.960	.962	.977	.970	.991	.960	.968	.797
High School Type (Baseline: Other)										
Humanity	.951	.952	.949	.955	.979	.970	.965	.986	.953	.962
Science	.894	.905	.926	.927	.951	.889	.916	.926	.045	.000
Tech	.790	.807	.790	.808	.819	.619	.751	.793	.003	.000
HS Grade	.955	.958	.972	.978	.971	.981	.987	.990	.984	.986
Year (Baseline: 2004)										
2005	.932	.964	.954	.926	.973	.977	.976	.983	.861	.918
2006	.883	.918	.914	.909	.959	.934	.952	.970	.807	.884
University (Pisa)	.950	.916	.971	.983	.686	.097	.225	.300	.082	.000
Major in University (Baseline: Other)										
Humanity	.946	.899	.689	.797	.798	.932	.958	.990	.964	.946
Science	.894	.857	.660	.751	.783	.901	.929	.966	.911	.913
Social Science	.798	.821	.624	.713	.758	.864	.913	.953	.878	.858
Bio-Med	.728	.776	.604	.677	.736	.837	.889	.926	.839	.832
Tech	.632	.634	.565	.624	.699	.794	.863	.876	.719	.453

TABLE 3

Posterior median and 95% credible intervals of principal strata proportion and super-population and finite-sample causal effects on dropout for always-applicants ( $\tau_{AA,s_0}$ ), compliant-applicants ( $\tau_{CA,s_0}$ ), and their union ( $\tau_{s_0}$ ), for the subpopulation within different bandwidths  $h$  around the threshold.

$h$	Population-average		Sample-average		Population-average at $s_0$	
	Median	95% CI	Median	95% CI	Median	95% CI
$h = 500$						
$\Pr(G_i = AA)$	.323	(.294; .355)	.322	(.309; .336)	.320	(.291; .352)
$\Pr(G_i = CA)$	.060	(.031; .105)	.041	(.021; .090)	.058	(.030; .094)
$\Pr(G_i = NA)$	.616	(.570; .650)	.637	(.590; .651)	.621	(.583; .654)
$\tau_{AA,s_0}$	-.153	(-.313; -.030)	-.152	(-.307; -.038)	-.154	(-.298; -.030)
$\tau_{CA,s_0}$	.045	(-.170; .497)	.074	(-.256; .545)	.039	(-.169; .474)
$\tau_{s_0}$	-.116	(-.253; -.005)	-.120	(-.265; -.009)	-.120	(-.245; -.012)
$h = 1\,000$						
$\Pr(G_i = AA)$	.336	(.312; .365)	.333	(.318; .354)	.335	(.311; .363)
$\Pr(G_i = CA)$	.043	(.002; .086)	.027	(.002; .075)	.043	(.001; .075)
$\Pr(G_i = NA)$	.623	(.584; .652)	.640	(.599; .645)	.625	(.594; .656)
$\tau_{AA,s_0}$	-.161	(-.273; -.052)	-.161	(-.270; -.057)	-.154	(-.259; -.052)
$\tau_{CA,s_0}$	.028	(-.745; .828)	.031	(-.778; .871)	.010	(-.918; .933)
$\tau_{s_0}$	-.132	(-.242; -.021)	-.139	(-.247; -.034)	-.128	(-.229; -.020)
$h = 1\,500$						
$\Pr(G_i = AA)$	.332	(.315; .349)	.332	(.326; .337)	.329	(.312; .346)
$\Pr(G_i = CA)$	.042	(.035; .077)	.027	(.020; .066)	.042	(.036; .062)
$\Pr(G_i = NA)$	.625	(.591; .642)	.642	(.605; .644)	.628	(.606; .646)
$\tau_{AA,s_0}$	-.183	(-.286; -.077)	-.187	(-.291; -.085)	-.153	(-.247; -.063)
$\tau_{CA,s_0}$	.010	(-.304; .797)	.011	(-.207; .928)	.000	(-.154; .951)
$\tau_{s_0}$	-.153	(-.256; -.040)	-.165	(-.266; -.057)	-.130	(-.217; -.019)

TABLE 4

Bayesian  $p$ -values of signal, noise and SNR under different  $h$  for the model used in the application to Italian university grants.

$h$	Principal strata	Signal	Noise	SNR
500	{CA}	.095	.630	.094
	{AA}	.254	.325	.254
	{AA,CA}	.338	.273	.370
1000	{CA}	.411	.425	.419
	{AA}	.400	.444	.473
	{AA,CA}	.493	.335	.518
1500	{CA}	.208	.444	.210
	{AA}	.372	.400	.261
	{AA,CA}	.455	.337	.470